


```
Db 10250 TGGACGCGCTGTAGTCCAGCTACTTGGAGGCTGTAGGAGGAGATCGCTTGAACCT 10309
QY 61 GGGAGACGGAGTTGCAGTACGACGACGATCGGTCTCACTCCAGCCTGGCGACAGAC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10310 GGGAGCGGAGGTTGCGGTGAGCGGAGATGGGCCACTGCCTCCAGCTGGCGACAGAG 10369
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 121 GTTCCGTTTCAAAAGAAAAAATAATTAATAAAAAAGAAATAAAA 165
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10370 CGACACTCTGTCTCAAAAAAACAACAAAAAACAACAAAAA 10414

RESULT 2
US-09-791-211-3
; Sequence 3, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 3
; LENGTH: 87543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 7421
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 7427
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 11609
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12605
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 12742
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29370
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29422
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29979
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29980
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 29981
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30136
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 30140
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31205
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31206
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 31592
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33095
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 33160
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34066
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 34072
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 36816
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 39020
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42164
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 42459
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46808
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46823
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 46826
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 47291
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 52786
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 52787
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 53384
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 54684
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59215
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59235
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 59242
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 63290
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 66614
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68660
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68697
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 68718
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
```



```

; PRIOR FILING DATE: 1998-03-16
;
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

```

Query Match 11.0%; Score 110.6; DB 4; Length 72604;
Best Local Similarity 79.4%; Pred. No. 1.9e-22;
Matches 131; Conservative 0; Mismatches 34; Indels 0;

QY	1	TGGCAGCGCCTGTAGTCCCACTACTCAGAGACTCAGGACAGAGAATCGTTGAACCC	60
Db	49134	TGGCAGCTGCCTGTAGTCCCACTACTCAGTGGCTGAGSCAGAAGATCGTTGAACCT	49075
QY	61	GGGACGACGGATTGCATGTAGTGCACCAAGATCGCGTCTACTGCACCTCCAGCTGCGCAGAC	120
Db	49074	AGGAGCGGAGGTTCTACTGAGCCAAAGATCGCACCTGCACCTCCAGCCTGAAGACAGAG	49015
QY	121	GTTCCGTTCAAAGAAAAAATAATATTTAAATAAGAAATAAAA	165
Db	49014	CGAGACTCTGTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAA	48970

```

RESULT 6
US-08-080-255-6
Sequence 6, Application US/08080255
Patent No. 5487970
GENERAL INFORMATION:
APPLICANT: Rowley, Janet D.
APPLICANT: Diaz, Manuel O.
TITLE OF INVENTION: COMPOSITION
TITLE OF INVENTION: DEPOSITING G
TITLE OF INVENTION: TRANSLOCATI
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Du
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

```

```

1 215. 77210
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: PatentIn Release #1.0, Version #1.25
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/08/080,255
16
17 FILING DATE: 19930617
18
19 CLASSIFICATION: 435
20
21 ATTORNEY/AGENT INFORMATION:
22
23 NAME: Parker, David L.
24
25 REGISTRATION NUMBER: 32,165
26
27 REFERENCE/DOCKET NUMBER: ARCD:072/PAR
28
29 TELECOMMUNICATION INFORMATION:
30
31 TELEPHONE: (512) 320-7200
32
33 TELEFAX: (512) 474-7577
34
35 INFORMATION FOR SEQ ID NO: 6:
36
37 SEQUENCE CHARACTERISTICS:
38
39 LENGTH: 8392 base pairs
40
41 TYPE: NUCLEIC ACID
42
43 STRANDEDNESS: single
44
45 TOPOLOGY: linear
46
47 MOLECULE TYPE: DNA (genomic)
48
49 US-08-080-255-6

```

Query Match	11.0%	Score 110.4;	DB 1;	Length 8392;
Best Local Similarity	82.7%;	Pred. No. 5.9e-23;		
Matches 139; Conservative	0;	Mismatches 26;	Indels 3;	Gaps 1;

Qy	1	TGGCAGCGGCTGTAGTCCCGACTACTCAGGAGACTCAGGCAGGAGAATCGCTTGAACCC	60
Dd	6269	TGGCAGCGGCTGTGATCCCGACTACTCAGGAGACTCAGGCTGGAGAATCGCTTGAACCC	6328
Qy	61	GGGAGCGGAGTTGTCAGTGAGCCCAAGATCGCTCATCGCACTCCAGCGTG---	117
Dd	6329	AGGAGCGGAGTTGTCAGTGAGTGCAGATGCCACCATTGCACCCCAGCTGGGCGACAG	6388
Qy	118	GACGTTCCGTTTCAAAAAGAAAAAATAATATTAATAAAAAAGAAATAAA	165
Dd	6389	GAGACTCCGCTCAAAAAATAATAATAATAATAATAATAATAATAATAATAA	6436

```

RESULT 7
US-08-465-713--6
Sequence 6, Application US/08465713
Patent No. 6121419
GENERAL INFORMATION:
APPLICANT: Rowley, Janet D.
APPLICANT: Diaz, Manuel O.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: DETECTING GENE REARRANGEMENTS AND
TITLE OF INVENTION: TRANSLOCATIONS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,713
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/080,255
FILING DATE: 17 JUNE 1993
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: ARCD:072/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 320-7200
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 8392 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-465-713--6

```

	Query Match	11.0%	Score 110.4;	DB 3;	Length 8392;
	Best Local Similarity	82.7%;	pred. No. 5.9e-23;		
	Matches 139;	Conservative 0;	Mismatches 26;	Indels 3;	Gaps 1;
QY	1	TGGCAGCGGCGTGTAGTCCCGACTACTCAGGAGACTCAGCAGAGAAATCGCTTGAACCC	60		
Db	6269	TGGCAGCGGCGTGTAGTCCCGACTACTCAGGAGACTCAGGAGAAATCGCTTGAACCC	6328		
QY	61	GGGAGAGCGGAGGTTGCAGTGCAGCAAGATCGCGTCACTGCAGTCCACGCGTG	117		

Qy	61	GGGAGACGGAGGTTGCAGTGTAGCGCAAGATCGCGTGCACATGCACATCCAGCCTGGCGACAGAC	120
Db	8495	AGGAGCAGAGGTTGCAGTGTAGCGGAGATCGTGGCACTGCACATCCAGCCTGGCAACAGAG	8554
Qy	121	GTTCGGTTTCAAAGAAAAAATAATATTAATAAAGAATAAAATCCGGCGCTGCCGGG	180
Db	8555	CGAGACTCCGCTCTCAAAAAAATAAAGAAAAAACAACAAAAAACAACAAACCAACCTTCA	8614
Qy	181	TGACATCAGTCTCTGTGCTTAATATGCGTCGCGCGCGCTACCGTCTCGGCAG	231
Db	8615	TGCGAAGATCTAGATTGTCTGCTGAATAACTGTGGATCTCGCTAGCCATAGCAAA	8665

```

RESULT 11
US-09-754-250-3
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

```

```

RESULT 12
US-09-488-856A-10/c
; Sequence 10, Application US/09488856A
; Patent NO. 6316259
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Robert McKay
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF GLYCOGEN SYNTHASE KINASE 3 ALPHA EXP
; FILE REFERENCE: RTS-0115
; CURRENT APPLICATION NUMBER: US/09/488,856A
; CURRENT FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 10
; LENGTH: 12394
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(397)
; NAME/KEY: CDS
; LOCATION: (2438)...(2625)
; NAME/KEY: CDS
; LOCATION: (5639)...(5722)
; NAME/KEY: CDS
; LOCATION: (5864)...(5974)
; NAME/KEY: CDS
; LOCATION: (7902)...(8032)
; NAME/KEY: CDS
; LOCATION: (8121)...(8227)
; NAME/KEY: CDS
; LOCATION: (9197)...(9294)
; NAME/KEY: CDS
; LOCATION: (9375)...(9470)
; NAME/KEY: CDS
; LOCATION: (9898)...(10084)
; NAME/KEY: CDS
; LOCATION: (10431)...(10523)
; NAME/KEY: CDS
; LOCATION: (11713)...(11786)
;
US-09-488-856A-10

Query Match          10.7%; Score 107; DB 4; Length 12394;
Best Local Similarity 81.4%; Pred. No. 8e-22;
Matches 136; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

Qy      1  TGGCAGCGGCTGTGAGTCCAGCTTACTCAGGAGACTGAGGCAGGAGAACTGCTTTGAACCC 60
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      4836 TGGCAGGCACTGTGATGCCAGCTACTTGGGAGGCTGAGGCAGGAGAAATGCGCTGAACCC 4777

Qy      61  GGGAGAGCGAGGTTGTCAGTGAGCCAAAGATCCGCTCACTGCACTCCAGCCT-GGGCGACAGA 119
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      4776 AGGAGGCGGAGCTTGCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGA 4717

Qy      120 CGTTCCGTTTTCAAAGAAAAAATAATATTATTAATAAGAAATAAAT 166
         |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      4716 CTCGTGCTCAAAAAATAATAATAATAATAATAATAATAATAAGATAAAT 4670

```

RESULT 13
US-08-709-923-1/c
; Sequence 1, Application US/08709923
; Patent No. 5831055
; GENERAL INFORMATION:
; APPLICANT: BIENKOWSKI, MICHAEL J.
; TITLE OF INVENTION: NOVEL KIDNEY ATP-DEPENDENT POTASSIUM
; CHANNELS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia & Upjohn Company
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: MI
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/709,923
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Woolton, Thomas A.
; REGISTRATION NUMBER: 35,004
; REFERENCE/DOCKET NUMBER: 6001.N CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616-833-7914

```

; TELEFAX: 616-833-6897
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2896 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-709-923-1

```

	Query Match	10.6%;	Score 106.6;	DB 2;	Length 2896;
	Best Local Similarity	76.9%;	Pred. No. 4.4e-22;		
	Matches 130;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps
QY	1	TGGCAGCGGCTCTAGTCCAGCTACTCAGGAGACTGAGCGAGGAGAAATCGCTTGAACCC	60		
Db	169	TGGCATGCGGCTCTAGTCCAGCTACTCAGGAGCTGAGCGAGGAGAAATCGCTTGAACCC	110		
QY	61	GGGAGACGAGGTTGCGAGTGAGCCAAGATCGGCTCACTGCACCTCCAGCCTGGCGACAGAC	120		
Db	109	GGGAGACAGAGTTGCGAGTGAGCGGACATCGCGCCACTGCACCTCCAGCCTGGGGAAACA	50		
QY	121	GTTCGGTTTCAAAAGAAAAAATAATTAATAAAAAGAAATAAATTCG	169		
Db	49	GGAAATTCACACTCAAGAAAAAGAGAGAGAGATTAAGATCAGATTCAG	1		

RESULT 14
US-08-724-394A-20/c
; Sequence 20, Application US/08724394A

```

: GENERAL INFORMATION:
:
: APPLICANT: Feder, John N.
: APPLICANT: Kronmal, Gregory S.
: APPLICANT: Lauer, Peter M.
: APPLICANT: Ruddy, David A.
: APPLICANT: Thomas, Winston
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
:
: TITLE OF INVENTION: Megabase T
:
: TITLE OF INVENTION: Sequences
:
: NUMBER OF SEQUENCES: 31
:
: CORRESPONDENCE ADDRESS:

```

```

; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc:feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note="HLA-H.CONTIG"
; US-08-724-394A-20

```

Query Match	10.6%	Score 106.6	DB 2	Length 246240
Best Local Similarity	81.8%	Pred. No. 6.3e-21		
Matches 135	Conservative 0	Mismatches 29	Indels 1	Gaps 1
Qy 1	TGGCAGCGGCTGTAGTCCCGACTACTCAGCAGACTGTAGCGAGGAGAAATCGCTTGAACCC	60		
Db 172447	TGGCACATGCTGTAGTCCCGACTACTGGGAGGCTGTAGCGAGGAGAAATGCTTGAATC	172388		
Qy 61	GGGAGAGCGAGGTTGTCAGTGTAGCCCAAGATCGCTGCTACCTGCAGCTTGCAGCAGAC	120		
Db 172387	GGGAGCGGAGGTTGCACGAGCCGAGATTGGCCACTGCACTCAGCCTGGTGACAG-	172329		
Qy 121	GTTCGGTTTCAAAAGAAAAAATAATATTAATAAAAAAGATAAAA	165		
Db 172328	GTGAGATCCGCTCAAAAATAATAATAATAATAATAATAATAATA	172284		

RESULT 15
US-08-724-394A-21/c
; Sequence 21, Application US/08724394A

```

: AGENT NO. 50422
:
: GENERAL INFORMATION:
:
: APPLICANT: Feder, John N.
: APPLICANT: Kronmal, Gregory S.
: APPLICANT: Lauer, Peter M.
: APPLICANT: Ruddy, David A.
: APPLICANT: Thomas, Winston
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
:
: TITLE OF INVENTION: Megabase Transcript Map: No. 587223761
:
: TITLE OF INVENTION: Sequences and Antibodies Thereto
:
: NUMBER OF SEQUENCES: 31
:
: CORRESPONDENCE ADDRESS:
:

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.30
OPENING SYSTEM: PC DOS/FAT DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:

```

```
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match      10.6%; Score 106.6; DB 2; Length 246240;
Best Local Similarity 81.8%; Pred. No. 6.3e-21;
Matches 135; Conservative 0; Mismatches 29; Indels 1; Gaps 1;

Qy 1 TGGCAGCGCCCTGTAGTCCCACTACTCAGGAGACTGAGCGAGGAGAAATCGCTTGAACCC 60
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172447 TGGCACATGCCTGTAGTCCCACTACTGCGGAGGCTGAGGCGAGGAAATTCGTTGAACCTC 172388

Qy 61 GGGAGCGGAGGTTGCAGTGAGCCCAAGATCCGCTCACTGCACCTCAGCCTGGCGACAGAC 120
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172387 GGGAGCGGAGGTTGCAGTGAGCCCAAGATTCGCCCACTGCACCTCAGCCTGGTGACAGA- 172329

Qy 121 GTTCCGTTTCAAAAGAAAAAATAATATTAATAAAAGATAAAA 165
      || || ||||| ||||| ||||| || |||||
Db 172328 GTGACACTCCGCTCAAAAATAATATAATAATAATAATAATA 172284

Search completed: December 24, 2002, 23:36:36
Job time : 489.673 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 21:54:28 ; Search time 32.3575 Seconds
(without alignments)
12277.043 Million cell updates/sec

Title: US-09-708-724A-3_COPY_99000_100000

Perfect score: 1001

Sequence: 1 tggcagcgctgtagctcc.....ttttatccaccatcaactaa 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 356696 seqs, 198428768 residues

Total number of hits satisfying chosen parameters: 713392

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c	1	117.4	11.7	17792	US-09-764-869-1599
	2	115	11.5	601	US-09-820-003A-28
	3	115	11.5	46050	US-09-820-003A-3
	4	113.6	11.3	30350	US-10-118-328-3
	5	113.4	11.3	35414	US-09-836-607-45
c	6	113.4	11.3	397658	US-09-813-320-3
	7	112.4	11.2	2147	US-09-764-847-1738
	8	112.2	11.2	14485	US-09-876-216-3
	9	112.2	11.2	95014	US-09-880-107-3428
	10	111.2	11.1	2547	US-09-764-877-2582
c	11	111.2	11.1	1503841	US-09-946-807-1
	12	111.2	11.1	1503841	US-09-795-668-1
	13	111.2	11.1	1503841	US-09-795-686-1
	14	110.6	11.0	15849	US-09-880-107-2362
	15	110.6	11.0	32154	US-09-764-877-3433
c	16	110.4	11.0	406	US-09-764-868-1475
	17	110.4	11.0	2294	US-10-052-586-283
	18	110.4	11.0	31208	US-09-852-067-3
	19	110.4	11.0	198285	US-09-880-107-3814

c	20	110.2	11.0	145831	10	US-09-969-708-79	Sequence 79, Appl
c	21	110.2	11.0	145831	10	US-09-954-456-2116	Sequence 2116, Ap
c	22	110	11.0	1400	10	US-09-263-959-295	Sequence 295, App
c	23	110	11.0	29695	10	US-09-752-820A-3	Sequence 3, Appli
c	24	110	11.0	29695	10	US-09-813-319A-3	Sequence 3, Appli
c	25	110	11.0	684973	10	US-09-263-959-1	Sequence 1, Appli
c	26	109.8	11.0	57130	10	US-09-835-081-3	Sequence 3, Appli
c	27	109.4	10.9	11991	10	US-09-764-877-2942	Sequence 2942, Ap
c	28	109	10.9	213	9	US-09-860-670-177	Sequence 177, App
c	29	109	10.9	21724	10	US-09-764-864-1603	Sequence 1603, Ap
c	30	109	10.9	21727	10	US-09-764-864-1604	Sequence 1604, Ap
c	31	109	10.9	110096	10	US-09-880-107-1542	Sequence 1542, Ap
c	32	108.6	10.8	248	10	US-09-764-847-1741	Sequence 1741, Ap
c	33	108.6	10.8	3807	10	US-09-764-877-2864	Sequence 2864, Ap
c	34	108.4	10.8	13058	10	US-09-764-846-303	Sequence 303, App
c	35	108.4	10.8	23433	10	US-09-927-091-7	Sequence 7, Appli
c	36	108.4	10.8	30625	10	US-09-927-091-5	Sequence 5, Appli
c	37	108.4	10.8	30676	10	US-09-927-091-8	Sequence 8, Appli
c	38	108.4	10.8	32185	10	US-09-764-877-3171	Sequence 3171, Ap
c	39	108.4	10.8	65608	9	US-09-954-531-180	Sequence 180, App
c	40	108.4	10.8	65608	10	US-09-962-436-292	Sequence 292, App
c	41	108.4	10.8	65608	10	US-09-962-832-119	Sequence 119, App
c	42	108.2	10.8	1807	10	US-09-880-107-2132	Sequence 2132, Ap
c	43	108	10.8	7537	10	US-09-764-869-1735	Sequence 1735, Ap
c	44	108	10.8	145831	10	US-09-969-708-79	Sequence 79, Appl
c	45	108	10.8	145831	10	US-09-954-456-2116	Sequence 2116, Ap

ALIGNMENTS

RESULT 1

US-09-764-869-1599/c
; Sequence 1599, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1599
; LENGTH: 17792
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1599

Query Match 11.7%; Score 117.4; DB 10; Length 17792;
Best Local Similarity 81.1%; Pred. No. 8e+22; 31; Indels 4; Gaps 1;
Matches 150; Conservative 0; Mismatches

QY	1	TGGCAGCGCGCTGTAGTCCAGCTACTCAGGAGACTGAGGAGAGAAATCGCTTGAACCC	60
DB	8122	TGGCAGCGCGCTGTAGTCCAGCTACTCAGGAGCTGAGGAGAGAAATCGCTTGAACCT	8063
QY	61	GGGAGACGAGGTTGCAGTGAGCCCAAGATCGCGTCACTGCTCCAGCTCGGCGACA---	117
DB	8062	GGGAGCGGAGGTTGCAGTGAGCTGAGATTCACCACTCCAGCTCGGCGACAAG	8003
QY	118	-GACGTTCCGTTTCAAGAGAAAAAATAATATATAAAGAAATAAATTCGGCGCTGC	176
DB	8002	CGAGACTCGCTCTCAAAAAAATAATAATAATAATAATAATAATAATAATAATAATA	7943
QY	177	CGCGT 181	
DB	7942	ATGGT 7938	
RESULT 2			
US-09-820-003A-28			

Matches 143; Conservative 0; Mismatches 26; Indels 4; Gaps 1;

QY 2 GGCACGCGCCTGTAGTCCAGCTACTCAGGAGACTGAGGAGAGAAATCGCTTGAACCG 61
|||||
Db 25521 GGCATGCGCCTGTAGTCCAGCTACTCCGAGGCTGAGGAGAGAAATCGCTTGAACCTG 25462

QY 62 GGAGACGGAGGTTGCAGTGAGCCAGCAAGATCGCGTCACTGCACCTCCAGCCTGGCGACA---- 117
|||||
Db 25461 GGAGGAGAGAGTTGCACTGAGCGAGATCGTGCACCTGCACCTCCAGCTTGGTGACAGAGC 25402

QY 118 GAGCTTCCGTTTCAAAGAAAAAATAATTAATAAAAAAGATAAAATCCGG 170
|||||
Db 25401 GAGACTCTGTCTCAAAAAATAAAATAAAATAAGAAAAATAATAATCCGG 25349

RESULT 6
US-09-813-320-3
; Sequence 3, Application US/09813320
; Patent No. US20020142378A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, Hongyu et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001172
; CURRENT APPLICATION NUMBER: US/09/813,320
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 397658
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...((397658)
; OTHER INFORMATION: n = A,T,C or G
US-09-813-320-3

Query Match 11.3%; Score 113.4; DB 10; Length 397658;
Best Local Similarity 78.9%; Pred. No. 5.2e-20;
Matches 135; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 TGCACGCGCCTGTAGTCCAGCTACTCAGGAGACTGAGGAGAGAAATCGCTTGAACCG 60
|||||
Db 195582 TGCACGTGCTGTAGTCCAGCTACTCAGGAGCTGAGGAGAGAAATCACTTGAACCC 195641

QY 61 GGAGACGGAGGTTGCAGTGAGCCAGATGCGGTCACTGCACCTCCAGCCTGGCGACAGAC 120
|||||
Db 195642 GGGAGCGGAGGTTGCAGTGAGCCAGATCACTGCACCTCCAGCCTGGTGACAGAG 195701

QY 121 GTTCCGTTTCAAAGAAAAAATAATTAATAAAAAAGATAAAATCCCG 171
|||||
Db 195702 CAAATTCCTTTAAAAAATAAAATGCTTCAATAATAATAAACTCAATCAAGC 195752

RESULT 7
US-09-764-847-1738
; Sequence 1738, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1738
; LENGTH: 2147
; TYPE: DNA
; ORGANISM: Homo sapiens

Result No.	Score	Query			DB	ID	Description
		Match	Length				
C 1	77.2	7.7	368004	10	US-09-949-654-3	Sequence 3, Appli	
C 2	76	7.6	303250	10	US-09-862-832-154	Sequence 154, App	
C 3	73.2	7.3	326014	10	US-09-731-231A-3	Sequence 3, Appli	
C 4	70.8	7.1	465237	10	US-09-933-267A-1	Sequence 1, Appli	
C 5	70.2	7.0	678	10	US-09-811-284-116	Sequence 116, App	
C 6	70	7.0	88191	10	US-09-799-799-3	Sequence 3, Appli	
C 7	69.4	6.9	560	10	US-09-967-701-1522	Sequence 1522, Ap	
C 8	66.6	6.7	197997	10	US-09-832-246-3	Sequence 3, Appli	
C 9	65.8	6.6	32038	10	US-09-764-878-292	Sequence 292, App	
C 10	65	6.5	1400	10	US-09-263-959-299	Sequence 299, App	
C 11	65	6.5	684973	10	US-09-263-959-1	Sequence 1, Appli	
C 12	64.8	6.5	4409	10	US-09-764-869-2014	Sequence 2014, Ap	
C 13	64.8	6.5	4494	10	US-09-764-869-2012	Sequence 2012, Ap	
C 14	64.8	6.5	4495	10	US-09-764-869-2011	Sequence 2011, Ap	
C 15	64.2	6.4	45845	10	US-09-927-091-6	Sequence 6, Appli	
C 16	63	6.3	465	10	US-09-864-761-5381	Sequence 5381, Ap	
C 17	63	6.3	368004	10	US-09-949-654-3	Sequence 3, Appli	
C 18	62.6	6.3	503	10	US-09-783-590-7878	Sequence 7878, Ap	
C 19	61.4	6.1	465237	10	US-09-933-267A-1	Sequence 1, Appli	

Qy 411 CACACCTTGATGTTGGACTT 430
| | | | |
Db 32924 GCCTCCTTGATCTCAGACTT 32905

```

RESULT 2
US-09-962-832-154
; Sequence 154, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962.832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
; LENGTH: 302250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-154

```

Query Match	7.6%	Score 76	DB 10	Length 302250
Best Local Similarity	65.6%	Pred. No. 9.3e-09		
Matches 143	Conservative 0	Mismatches 70	Indels 5	Gaps 2
QY 215	TGTCGCTAGTTACAATCAAAAAGTTTTAAAATGAGGACCTTAGGTGGTGCTCTAATCCAA	274		
Db 113304	TGGGCGCTTTAAATAGGTAATTAGATAGAAATGAGCGCTTAGGTGGTGCTCTGATCCAA	113363		
QY 275	TCTAAGTGATGT-CTCCATGAAGAGGAAATPAGGATACAAATGTGCACACAGAGAGAA	333		
Db 113364	TATGACTGGTGTCATGTGTAAAAAAGGAGATTTCAGACACAGACTTGTGCAGAGGGAGAA-	113422		
QY 334	TGGCCACATGAGGACACAATGACGAATGTGGCTACTTACAAGCCTTAGGAGAGAGCCCTCG	393		
Db 113423	--CCATGTGCGGAGCGAGGAGAAATGGGCCATCTACAGCCAAAGGACAGAGCCCTCAG	113479		
QY 394	AGAAAACACACCCCTACCCACACCTTGATGTGGACTTC	431		
Db 113480	AATGAACCAACCCCTGCCACACACTTGGTCTCCAATTC	113517		

```

RESULT 3
US-09-731-231A-3/c
; Sequence 3, Application US/09731231A
; Patent No. US20020082189A1
; GENERAL INFORMATION:
; APPLICANT: GUGGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001007
; CURRENT APPLICATION NUMBER: US/09/731,231A
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 326014
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(326014)
; OTHER INFORMATION: n = A,T,C or G
; US-09-731-231A-3

```

Query Match	7.3%;	Score 73.2;	DB 10;	Length 326014;
Best Local Similarity	68.0%;	Pred. No. 4.8e-08;		
Matches 132;	Conservative 0;	Mismatches 58;	Indels 4;	Gaps 2;
QY	240	TTAAATGAGGACCTTAGGTGGGTCTTAATCCAAATCTAAGTGTATGTCCTCAT---	GAAG	296
Db	37807	TTAAATGAGGCCACGAGAGTGGCCCTAATCTTATATAGTGTGTCTTATAAGGAGA	37748	
QY	297	GAGGAATAAGGATACAAATGTGCACACAGAGAGAAATGCCACATGAGGACACAATGAG	356	
Db	37747	GGGAGAGACCAGGACACGCTTGGCACAAACACAGAAA--GGCCATGTGAAAACAGAGCAAG	37689	
QY	357	AATGTGCTACTTACAAGCCTTAGGAGAGAGCGCTCCGAGAAAAACACACCTTACCCACACC	416	
Db	37688	GAGTGGCCATCTACAAGCAAGGAGAGAGCGCTTCAGAAAAAATCACTCTGTGGCACC	37629	
QY	417	TTGATGTTGGACTT	430	
Db	37628	TTATCTTGGACTT	37615	
RESULT 4				
US-09-933-267A-1/c				
; Sequence 1, Application US/09933267A				
; Patent No. US20020123095A1				
; GENERAL INFORMATION:				
; APPLICANT: Kalush, Francis et al.				
; TITLE OF INVENTION: Estrogen receptor alpha variants and				
; METHOD OF INVENTION: methods of detection thereof				
; FILE REFERENCE: C1000258C14				
; CURRENT APPLICATION NUMBER: US/09/933,267A				
; CURRENT FILING DATE: 2001-08-21				
; PRIOR APPLICATION NUMBER: 60/160626				
; PRIOR FILING DATE: 1999-10-20				
; PRIOR APPLICATION NUMBER: 60/183756				
; PRIOR FILING DATE: 2000-02-22				
; PRIOR APPLICATION NUMBER: 09/692414				
; PRIOR FILING DATE: 2000-10-20				
; PRIOR APPLICATION NUMBER: 09/768184				
; PRIOR FILING DATE: 2001-01-24				
; PRIOR APPLICATION NUMBER: 09/804076				
; PRIOR FILING DATE: 2001-03-13				
; PRIOR APPLICATION NUMBER: 09/826314				
; PRIOR FILING DATE: 2001-04-05				
; NUMBER OF SEQ ID NOS: 2				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 1				
; LENGTH: 465237				
; TYPE: DNA				
; ORGANISM: human				
US-09-933-267A-1				
Query Match	7.1%;	Score 70.8;	DB 10;	Length 465237;
Best Local Similarity	63.4%;	Pred. No. 2.2e-07;		
Matches 128;	Conservative 0;	Mismatches 67;	Indels 7;	Gaps 1;
QY	240	TTAAATGAGGACCTTAGGTGGGTCTTAATCCAAATCTAAGTGTATGTCCTCATGAAAGAG	299	
Db	246245	TTCAATGAGGTCCTGTGTGTGGGCCCTTAATCCGATATGACTGGTGTCTTCGTAAGAAGA	246186	
QY	300	GAATAAGGATACAAATGTGCACACAGAGAGAAATGCCACATGAGGACACAATGAGAAAT	359	
Db	246185	AGAGAAG-----CTGTGGATCCCGAGAAAGGCCATGTGGGACACAGAGAGAAG	246133	
QY	360	GTGGCTACTTACAAGCCTTAGGAGAGAGCGCTCCGAGAAAAACACACCTTACCCACACCTTG	419	
Db	246132	GCAGCCATCTTCAGACCAAGGAGAGAGCGCTCGGGAAAAATCAACCTGCCAACACCTTA	246073	
QY	420	ATGTTGGACTTCATCTCTGTAGA	441	
Db	246072	GTCTCAGACTTTCAGCTCCAGA	246051	

```

RESULT 5
US-09-811-284-116/c
; Sequence 116, Application US/09811284
; Patent No. US20020058306A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; TITLE OF INVENTION: No. US20020058306alel G Protein-Coupled Receptors
; FILE REFERENCE: 00167051
; CURRENT APPLICATION NUMBER: US/09/811,284
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,907
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,918
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,960
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,917
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/192,945
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,916
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,923
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,933
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,830
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,234
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,155
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,935
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 116
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-811-284-116

Query Match      7.0%; Score 70.2; DB 10; Length 678;
Best Local Similarity 67.4%; Pred. No. 2e-08;
Matches 130; Conservative 0; Mismatches 58; Indels 5; Gaps 2;

Qy 240 TTAATAAGGAGGACCTTAGGGTGGTCTTAATCCCAATCTAAGTGATGCTCTCCATGA-AAGA 298
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 366 TTAATAAGGAGTCAATTAAGATGGCCCTGATGGAGCATAAACAGGTGCTCTTTATAAGAGA 307

Qy 299 GGAATAAGGATACAAATGTCCACACAGAGAGAAATGCCACATGAGGACACAATGAGAA 358
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 306 GGAGATTAGGACACAAA-----CACACACAGAGGGACGCCATGTGAAGACACAGCAAGAA 251

Qy 359 TGTGGCTACTTACAGCCTTAGGAGAGGCGCTCCGAGAAACACACACCTTACCCACACTT 418
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 250 GACGTCTCTGTGAGCCAGGAGAGAGGCGCTCAGAGAAACCAAAACCTGCCGACACCTT 191

Qy 419 GATGTTGGACTTC 431
|||||  |||  |||  |||
Db 190 GATCATGAACATC 178

RESULT 6
US-09-799-799-3
; Sequence 3, Application US/09799799
; Patent No. US20020132291A1
; GENERAL INFORMATION:
; APPLICANT: Ye, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

```

```

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001157
; CURRENT APPLICATION NUMBER: US/09/799,799
; CURRENT FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 88191
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(88191)
; OTHER INFORMATION: n = A,T,C or G
US-09-799-799-3

Query Match      7.0%; Score 70; DB 10; Length 88191;
Best Local Similarity 62.6%; Pred. No. 1.8e-07;
Matches 134; Conservative 0; Mismatches 65; Indels 15; Gaps 1;

Qy 218 GTCAGTTACAATCAAAAAGTTTAAATAGAGACCTTAGGGTGGTCTTAATCCCAATCT 277
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7364 GGCCTTTAAGAGAGGTAAACAAGTTAAATAGAGGTGCTGAGAGTGGTCTGTCTCAT 7423

Qy 278 AAGTGATGTCTCCATGAAGAGAGAAATAAGGATACAAATGTCCACACAGAGAGAAATGCC 337
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7424 GACTGGTGTCCAAAAAGAGAG-----CTGAGGACTCAGAGGGGACAAAC 7468

Qy 338 CACATGAGGACACAATGAGAAATGTGGCTACTTACAAGCTAGGAGAGAGGCGCTCGAGAA 397
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7469 CACGTGAGGACACAAGGATATGATGGTCTGCTTACAAGCAGGAGAGAGGCGCTCAGGAGA 7528

Qy 398 AACACACCCCTACCCACACCTTGATGTTGGACTTC 431
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7529 AACCAACCCCTGCCAACACCTTGATCTCAGAGTTC 7562

RESULT 7
US-09-867-701-1522/c
; Sequence 1522, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1522
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(560)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-1522

Query Match      6.9%; Score 69.4; DB 10; Length 560;
Best Local Similarity 67.5%; Pred. No. 2.9e-08;
Matches 129; Conservative 0; Mismatches 56; Indels 6; Gaps 2;

Qy 240 TTAATAAGGAGCCTTAGGGTGGTCTTAATCCCAATCTAAGTGATGCTCTCCATGAAGAG 299
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 202 TTAATAAGGAGCCATTTGGGTAGGCTTTAATCCAATATGACT--AGTATCCTTTATAAGAG 145

Qy 300 GAAATAAGGATACAAATGTGCACACAGAGAGAAATGCCACATGAGGACACAAATGAGAA 359
|||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 144 GAGATTAGAACACACAGACTAG----TATAGGGGAAAGACCACCATGTGAAGACACAGGGAGG 89

```


Db 158 GTGTCCTTATC-----AGGAGGTAGGACACAGACATGCAAGAGGGGAAGACAGAAAGAA 223

Qy 343 GAGGACACAATGAGAAATGGTCTACTTACAAAGCCTAGGAGAGAGGCGCTCCGAGAAAAACAC 402

Db 224 GGTATAGAAAAAATAATTTGTATCTACAAGCCAAAGAGAGAGGCGCTCAGAGAAACCA 283

Qy 403 ACCCTAC 409

Db 284 ACTCAAC 290

RESULT 11

US-09-263-959-1

; Sequence 1, Application US/09263959

; Patent No. US20020150891A1

GENERAL INFORMATION:

APPLICANT: Rowen, Lee

APPLICANT: Rowen, Ben F.

TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH UTI

NUMBER OF SEQUENCES: 1279

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: US

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/263,959

FILING DATE: 05-MAR-1999

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: McMasters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 920010.426C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 684973 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-263-959-1

Query Match 6.5%; Score 65; DB 10; Length 684973;

Best Local Similarity 63.1%; Pred. No. 7.5e-06;

Matches 118; Conservative 0; Mismatches 65; Indels 4; Gaps 1;

Qy 223 TTCAATCAAAAAAGTTTAAATGAGGACCTTAGGGTGGTCTTAATCCCAATCTTAAGTG 282

Db 544708 TTACATCGATAATTAAGTTAAATTAAGTTCATTAGGTGGGCCCTAATTCATATGACTG 544767

Qy 283 ATGCTCTCAATGAAGAGAAATAAGGATPACAAATGTGCACACAGAGAGAAATGGCCACAT 342

Db 544768 GTGTCCTTATC-----AGGAGGTAGGACACAGACATGCAAGAGGGGAAGACAGAAAGAA 544823

Qy 343 GAGGACACAATGAGAAATGGTCTACTTACAAAGCCTAGGAGAGAGGCGCTCCGAGAAAAACAC 402

Db 544824 GGTATAGAAAAAATAATTTGTATCTACAAGCCAAAGAGAGAGGCGCTCAGAGAAACCA 544883

Qy 403 ACCCTAC 409

Db 544884 ACTCAAC 544890

RESULT 12

US-09-764-869-2014/c

; Sequence 2014, Application US/09764869

; Patent No. US20020061521A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC007

CURRENT APPLICATION NUMBER: US/09/764,869

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 2442

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2014

LENGTH: 4409

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-869-2014

Query Match 6.5%; Score 64.8; DB 10; Length 4409;

Best Local Similarity 63.0%; Pred. No. 1e-06;

Matches 121; Conservative 0; Mismatches 62; Indels 9; Gaps 1;

Qy 240 TTAATAAGGAGGACCTTAGGTGGTCTTAATCCCAATCTTAAGTGATGTCTCCATGAAGAG 299

Db 3801 TTAATAAGTGTCTACTAGGTGGGCCCAATCCCAAGATGACTGGTGTCTCTTTTG----- 3748

Qy 300 GAAATAAGGATACAAATGTGCACACAGAGAGAAATGGCCACATGAGGACACAAATGAGAAT 359

Db 3747 ---GGAACAGGAGACTAGTGCACACAGAGAGAAATGGCCACATGAGGACACACTGGACAG 3691

Qy 360 GTGGCTACTTACAAGCCTTAGGAGAGAGGCGCTCCGAGAAAAACACACCCCTACCACACCTTG 419

Db 3690 ACGCCACCTGCAAGCCAAAGGAGGGGCGCTCAGAAGACATCAAACTTAAGGATAACTTG 3631

Qy 420 ATGTTGGACTTC 431

Db 3630 ATCTTGCACTTC 3619

RESULT 13

US-09-764-869-2012/c

; Sequence 2012, Application US/09764869

; Patent No. US20020061521A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC007

CURRENT APPLICATION NUMBER: US/09/764,869

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 2442

SOFTWARE: PatentIn Ver.2.0

SEQ ID NO 2012

LENGTH: 4494

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-869-2012

Query Match 6.5%; Score 64.8; DB 10; Length 4494;

Best Local Similarity 63.0%; Pred. No. 1e-06;

Matches 121; Conservative 0; Mismatches 62; Indels 9; Gaps 1;

Qy 240 TTAATAAGGAGGACCTTAGGTGGTCTTAATCCCAATCTTAAGTGATGTCTCCATGAAGAG 299

Db 3801 TTAATAAGTGTCTACTAGGTGGGCCCAATCCCAAGATGACTGGTGTCTCTTTTG----- 3748

Qy 300 GAAATAAGGATACAAATGTGCACACAGAGAGAAATGGCCACATGAGGACACAAATGAGAAT 359

Db 3747 ---GGAACAGGAGACTAGTGCACACAGAGAGAAATGGCCACATGAGGACACACTGGACAG 3691

Qy 360 GTGGCTACTTACAAGCCTTAGGAGAGAGGCGCTCCGAGAAAAACACACCCCTACCACACCTTG 419

Db 3690 ACGGCCACCTGCAAGCCAAAGGGAGGGCGCTCAGAAGACATCAAACTTAAGGATAACTTGG 3631

Qy 420 ATGTTGGACTTC 431
|| ||| | |||

Db 3630 ATCTTGCACTTC 3619

RESULT 14
US-09-764-869-2011/c
; Sequence 2011, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2011
; LENGTH: 4495
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-2011

Query Match 6.5%; Score 64.8; DB 10; Length 4495;
Best Local Similarity 63.0%; Pred. No. 1e-06; Mismatches 9; Gaps 1;
Matches 121; Conservative 0;

Qy 240 TTAATAATGAGGACCTTAGGTGGTGCCTTAATCCATCTAAGTGATGTCCTCCATGAAAGAG 299
||||||| | ||||||| | ||||||| | ||||||| | ||||||| |

Db 3802 TTAATAATGAGTTCACCTAGGTGGGCCCCCAATCCAAAGATGACTGGTCTCTTTTG----- 3749

Qy 300 GAAATAAGGATACAAATGTCACACAGACAGAGAATGGCCACATGAGGACACAAATGAGAT 359
|| | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 3748 ---GGAACAGAGGAGACTAGGTCTCACACAGAGAAAAAGGCTGCGTGAGGACACACTGGACAG 3692

Qy 360 GTGGCTACTTACAAGCTAGGAGAGAGGCGCTCCGAGAAAAACACACCCCTACCCACACCTTG 419
|| | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 3691 ACGGCCACCTGCAAGCCAAAGGGAGGGCGCTCAGAAGACATCAAACTTAAGGATAACTTGG 3632

Qy 420 ATGTTGGACTTC 431
|| ||| | |||

Db 3631 ATCTTGCACTTC 3620

RESULT 15
US-09-927-091-6/c
; Sequence 6, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOTT, STEVE
; APPLICANT: CHANDLER, DAWN
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
; FILE REFERENCE: UTSC:651US
; CURRENT APPLICATION NUMBER: US/09/927,091
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 45845
; TYPE: DNA
; ORGANISM: Human
US-09-927-091-6

Query Match 6.4%; Score 64.2; DB 10; Length 45845;
Best Local Similarity 59.7%; Pred. No. 3.8e-06;

Matches 126; Conservative 0; Mismatches 83; Indels 2; Gaps 1;

Qy 221 AGTTACAATCAAAAAAGTTTTAAATGAGGACCTTAGGTGGTGGTCTTAATCCAATCTAAG 280
|| || | | |||| | || | | | | | | | | | | | | | | | | | |

Db 3730 AGATAGGATATTTAAAGAGATAATCAGGTTAAATTAGGTGGCCCTAATCTACTATGAC 3671

Qy 281 TGATGCTCCATGAAAGAGGAAATAAGGATACAAATGTCACACAGAGAGAAATGGCCAC 340
|| |||| | | |||| | | |||| | | |||| | | |||| | | |||| | |

Db 3670 TGGTGCTCT-TTGTAAGAAGAGAGAATTTGGACACAGACAAGGTACAGAGGGAATATGAT 3613

Qy 341 ATGAGGACACAATGAGAATGTGGCTACTTACAAGCCTAGGAGAGAGGCTCCGAGAAAAAC 400
|| | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 3612 GTGAAGACACAGGGGAGAGACCGTCATCTACAAGACAAGGAGAGAGGCTCAGAGAGACAT 3553

Qy 401 ACACCTACCCACACCTTTGATGTTGGACTTC 431
|||| | | | | | | | | | | | | | | | | | | | | | | | |

Db 3552 CAACCTGCTGAAATCTTAATCTTGGATTC 3522

Search completed: December 25, 2002, 01:44:33
Job time : 2039.36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 16:24:42 ; Search time 2253.71 Seconds
(without alignments)
16077.036 Million cell updates/sec

Title: US-09-708-724A-1
Perfect score: 1245
Sequence: 1 atggagaccctggtcagtggt.....ccaggaaaggcggtgtagtag 1245

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_on:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pi:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_on:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pi:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rtd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	229.4	18.4	151049	9	AC018558	Homo sapi
2	229.4	18.4	191768	9	AC092357	Homo sapi
3	203.4	16.3	191014	9	AC113268	Papio cyn
4	192	15.4	153629	9	AC092325	Homo sapi
5	183.2	14.7	125133	9	HSBG8211	Human DNA
6	173.8	14.0	185664	9	AC093249	Homo sapi
7	173.8	14.0	207538	2	AC106886	Homo sapi
8	173	13.9	124047	9	AL713966	Human DNA
9	173	13.9	131234	9	HS17282	Human DNA
10	173	13.9	160623	9	AL670296	Human DNA
11	169	13.6	183165	9	AP000802	Homo sapi
12	169	13.6	184841	9	AC016902	Homo sapi
13	169	13.6	185569	2	AP003099	Homo sapi
14	168.8	13.6	7736	9	AF079797	Homo sapi
15	168.8	13.6	35715	9	AC020951	Homo sapi
16	168.8	13.6	178184	9	AC008770	Homo sapi
17	168.8	13.6	179272	2	AC069149	Homo sapi
18	168.8	13.6	179581	2	AC009397	Homo sapi
19	166.2	13.3	178053	2	AC021112	Homo sapi
20	165.6	13.3	184539	2	AC027192	Homo sapi
21	165.6	13.3	192087	2	AC079325	Homo sapi
22	165.6	13.3	201306	2	AC087390	Homo sapi
23	165.6	13.3	206143	2	AC069390	Homo sapi
24	164.6	13.2	36430	9	AC010506	Homo sapi
25	163.8	13.2	121287	9	AC010332	Homo sapi
26	163.8	13.2	184169	2	AC026349	Homo sapi
27	163.8	13.2	219553	2	HS312687	Homo sapi
28	163.6	13.1	187280	2	AC092674	Homo sapi
29	163.6	13.1	207962	2	AC015679	Homo sapi
30	162.2	13.0	160169	2	AC051664	Homo sapi
31	160.8	12.9	24020	9	HSJ269L11	Human DNA
32	160.8	12.9	130819	9	AL662842	Human DNA
33	160.6	12.9	48993	9	AC104527	Homo sapi
34	160.6	12.9	159849	2	AC021454	Homo sapi
35	160.6	12.9	163915	2	AC087451	Homo sapi
36	160.6	12.9	166867	9	AP003733	Homo sapi
37	160.6	12.9	174305	2	AC074126	Homo sapi
38	160.6	12.9	231247	9	AC008543	Homo sapi
39	160	12.9	143068	6	AX335952	Sequence
40	160	12.9	143068	9	HS095626	Homo sapien
41	160	12.9	220965	2	HS312688	Homo sapi
42	159	12.8	37392	9	AC005946	Homo sapi
43	159	12.8	143619	9	AC008812	Homo sapi
44	159	12.8	185437	9	AC098613	Homo sapi
45	159	12.8	190628	2	AC024700	Homo sapi

ALIGNMENTS

RESULT 1
AC018558
LOCUS
DEFINITION Homo sapiens chromosome 16 clone RP11-80F22, complete sequence.
ACCESSION AC018558
VERSION AC018558.5 GI:16596530
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 151049)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished

REFERENCE	2 (bases 1 to 151049)	DOE Joint Genome Institute.
AUTHORS	DOE Joint Genome Institute.	
TITLE	Sequencing of Human Chromosome 16	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 191768)	DOE Joint Genome Institute.
AUTHORS	DOE Joint Genome Institute.	
TITLE	Direct Submission	
JOURNAL	Submitted (03-JUL-2001)	Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	3 (bases 1 to 191768)	DOE Joint Genome Institute.
AUTHORS	DOE Joint Genome Institute.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-AUG-2001)	Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On or before Aug 15, 2001 this sequence version replaced gi:795566.	
	Sequence Quality Assessment:	
	This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.	
	All manually edited bases have been reduced to quality zero.	
	Quality levels above 40 are expected to have less than 1 error in 10,000 bp.	
	Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.	

	Sequence Quality Assessment:	
	This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.	
	All manually edited bases have been reduced to quality zero.	
	Quality levels above 40 are expected to have less than 1 error in 10,000 bp.	
	Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.	

FEATURES	Location/Qualifiers	
source	1..151049	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="16"	
	/clone="RP11-80F22"	
BASE COUNT	46827 a 30783 c 30749 g 42690 t	
ORIGIN		
	Query Match 18.4%; Score 229.4; DB 9; Length 151049;	
	Best Local Similarity 97.5%; Pred. No. 2.8e-47;	
	Matches 233; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	692 CCAGGAGCCATGTTCTGAATCTCAAATTTTGAAGAATCTTTGCCCCACCAACAC 751	
Db	133275 CGCAGGAGCCATGTTCTGAATCTCAAATTTTGAAGAATCTTTGCCCCACCAACAC 133334	
QY	752 CCAAGAAAAATAAACAGGAGGAGGAGGATGAAAATTTGGCGTCTACACCCCTCCAG 811	
Db	133335 CCAAGAAAAATAAACAGGAGGAGGAGGATGAAAATTTGGCGTCTACACCCCTCCAG 133394	
QY	812 TAGCAGAACCTGTACCATCTCTTCAGTAACAGAAATAGACCCCATCGCAAGAA 871	
Db	133395 TAGCAGAACCTGTACCATCTCTTCAGTAACAGAAATAGACCCCATCGCAAGAA 133454	
QY	872 TTCCGGCGGACTGTACCATAGCTGGAGAGCCCTTAGGACATTCATTTCT 930	
Db	133455 TTCTATGCTGTCTACCATAGCTGGAGAGCCCTTAGGACATTCATTTCT 133513	
RESULT 2		
AC092357	191768 bp DNA linear PRI 15-AUG-2001	
LOCUS	AC092357	
DEFINITION	Homo sapiens chromosome 16 clone RP11-332P24, complete sequence.	
ACCESSION	AC092357	
VERSION	AC092357.2	
KEYWORDS	HTG.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 191768)	

AUTHORS	DOE Joint Genome Institute.	
TITLE	Sequencing of Human Chromosome 16	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 191768)	
AUTHORS	DOE Joint Genome Institute.	
TITLE	Direct Submission	
JOURNAL	Submitted (03-JUL-2001)	Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	3 (bases 1 to 191768)	
AUTHORS	DOE Joint Genome Institute.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-AUG-2001)	Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	On or before Aug 15, 2001 this sequence version replaced gi:795566.	
	Sequence Quality Assessment:	
	This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.	
	All manually edited bases have been reduced to quality zero.	
	Quality levels above 40 are expected to have less than 1 error in 10,000 bp.	
	Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.	

	Sequence Quality Assessment:	
	This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.	
	All manually edited bases have been reduced to quality zero.	
	Quality levels above 40 are expected to have less than 1 error in 10,000 bp.	
	Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.	

FEATURES	Location/Qualifiers	
source	1..191768	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="16"	
	/clone="RP11-332P24"	
BASE COUNT	57149 a 39592 c 40150 g 54877 t	
ORIGIN		
	Query Match 18.4%; Score 229.4; DB 9; Length 191768;	
	Best Local Similarity 97.5%; Pred. No. 2.8e-47;	
	Matches 233; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
QY	692 CCAGGAGCCATGTTCTGAATCTCAAATTTTGAAGAATCTTTGCCCCACCAACAC 751	
Db	83020 CGCAGGAGCCATGTTCTGAATCTCAAATTTTGAAGAATCTTTGCCCCACCAACAC 83079	
QY	752 CCAAGAAAAATAAACAGGAGGAGGATGAAAATTTGGCGTCTACACCCCTCCAG 811	
Db	83080 CCAAGAAAAATAAACAGGAGGAGGATGAAAATTTGGCGTCTACACCCCTCCAG 83139	
QY	812 TAGCAGAACCTGTACCATCTCTTCAGTAACAGAAATAGACCCCATCGCAAGAA 871	
Db	83140 TAGCAGAACCTGTACCATCTCTTCAGTAACAGAAATAGACCCCATCGCAAGAA 83199	
QY	872 TTCCGGCGGACTGTACCATAGCTGGAGAGCCCTTAGGACATTCATTTCT 930	
Db	83200 TTCTATGCTGTCTACCATAGCTGGAGAGCCCTTAGGACATTCATTTCT 83258	
RESULT 3		
AC113268/c	191014 bp DNA linear PRI 14-AUG-2002	
LOCUS	AC113268/c	
DEFINITION	Papio cynocephalus anubis clone rp41-22m16, complete sequence.	
ACCESSION	AC113268	
VERSION	AC113268.8	
KEYWORDS	HTG.	
SOURCE	olive baboon.	

ORGANISM Papio cynocephalus anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.

REFERENCE 1 (bases 1 to 191014)
AUTHORS Zhou, L., Fu, Y., Wu, J., Eichler, E. and Roe, B. A.
TITLE Papio anubis BAC Clone rp41-22m16
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 191014)
AUTHORS Zhou, L., Fu, Y., Wu, J., Eichler, E. and Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2002) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 3 (bases 1 to 191014)
AUTHORS Zhou, L., Fu, Y., Wu, J., Eichler, E. and Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (06-AUG-2002) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

REFERENCE 4 (bases 1 to 191014)
AUTHORS Zhou, L., Fu, Y., Wu, J., Eichler, E. and Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-2002) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

COMMENT On Aug 4, 2002 this sequence version replaced gi:21104942.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code: UOKNOR

FEATURES
source Location/Qualifiers
1. 191014
/organism="Papio cynocephalus anubis"
/sub_species="anubis"
/db_xref="taxon:9555"
/clone_lib="rp41-22m16"
/clone_lib="RP41 - 41 Male (Olive) Baboon BAC Library"
BASE COUNT 54882 a 40134 c 39560 g 56438 t
ORIGIN
Query Match 16.3%; Score 203.4; DB 9; Length 191014;
Best Local Similarity 91.1%; Pred. No. 1.2e-40;
Matches 216; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 694 CAGAGCCATGTTCTGAATCTCAAAATTTGAAAGAAATCTTTTGTCCACCCACACACCC 753
|||||
Db 122104 CAGAGCCATGTTCTGAATCTCAAAATTTGAAAGAAATCTTTTGTCCACCCACACACCC 122045
|||||

QY 754 AAGAAAAATATAACAGGAGGAGGAGTGAATAATGGCGTCTACACACCCCTCCAGTA 813
|||||
Db 122044 GATGAATAATATAACAGGAGGAGGAGTGAATAATGGCGTCTACACACCCCTCCAGTA 121985
|||||

QY 814 GCAGAAACACCTGTACCATCTCTCTAGTAAACAGAAATAGAGACCCCTCCAGGAATTT 873
|||||
Db 121984 GCAGAAACACCTGTATCATCTCTCTAGTAAACAGAAATAGAGACCCCTCCAGGAATTT 121925
|||||

QY 874 CCGGGGACTGCTACCATAGCTGGAGAGCCCTTAGGACATTCACATTTCTT 930
|||||
Db 121924 ATATGCTTTGCTGCCATAGCTGGAGAGCCCTTAGGACATTCACATTTCTT 121868
|||||

RESULT 4
AC092325/c 153629 bp DNA linear PRI 02-NOV-2001
LOCUS Homo sapiens chromosome 16 clone RP11-14K3, complete sequence.
DEFINITION AC092325 AC011682
AC092325.2 GI:16596537
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 153629)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished

REFERENCE 2 (bases 1 to 153629)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 153629)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On or before Nov 2, 2001 this sequence version replaced gi:7329393,
gi:14589514.
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source Location/Qualifiers
1. 153629
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-14K3"
BASE COUNT 45079 a 31598 c 31052 g 45900 t
ORIGIN
Query Match 15.4%; Score 192; DB 9; Length 153629;
Best Local Similarity 100.0%; Pred. No. 9.4e-38;
Matches 192; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 CAGCAGAACCATGAGCCTAGCGGAGCGCGGTTCGGAAGCTCCCTCCGCAACG 1113
|||||
Db 129738 CAGCAGAACCATGAGCCTAGCGGAGCGCGGTTCGGAAGCTCCCTCCGCAACG 129679
|||||

QY 1114 GGCCTCTCTCAGACCGCTCGTGCCTCCGCTGCGGAGCTGACCGCAGCGAGTGA 1173
|||||
Db 129678 GGCCTCTCTCAGACCGCTCGTGCCTCCGCTGCGGAGCTGACCGCAGCGAGTGA 129619
|||||

QY 1174 GAGTTGCTCTGTTGGCGGTGACCCAGGAGGAGTGGAGCGCGGATCATCTCCAGGAAG 1233
|||||
Db 129618 GAGTTGCTCTGTTGGCGGTGACCCAGGAGGAGTGGAGCGCGGATCATCTCCAGGAAG 129559
|||||

QY 1234 CGGGCTGAGTAG 1245
|||||

Db 129558 CGGGCTGAGTAG 129547
|||||

RESULT 5
HSBG8211 125133 bp DNA linear PRI 16-DEC-2000
LOCUS Human DNA sequence from clone GS1-8211 on chromosome Xq26.3-27.3
DEFINITION Contains a pseudogene similar to Novel human gene mapping to
chromosome X, gene for LDOC1 protein, ESTs, STSs, GSSs and a Cpg


```
consensus"
55571..55864
/note="ALUSg1 repeat: matches 1..294 of consensus"
55866..56210
/note="match: GSS: Em:A0310669"
complement(join(56599..56669,57317..57592))
/note="match: STS: Em:AA412270"
56677..57126
/note="LTR46 repeat: matches 1..461 of consensus"
complement(57127..57499)
/note="match: GSS: Em:AQ390433"
59073..59927
/note="LIM1 repeat: matches -1389..-530 of consensus"
60584..60929
/note="LIM48 repeat: matches 5937..6289 of consensus"
63785..64322
/note="LTR25-internal repeat: matches 3797..4338 of
consensus"
64318..64584
/note="MER4-internal repeat: matches 2241..2494 of
consensus"
64589..64668
/note="LTR8 repeat: matches 157..253 of consensus"
64674..64966
/note="MERS1-internal repeat: matches 3973..4273 of
consensus"
64903..66811
/note="MER41-internal repeat: matches 1826..3447 of
consensus"
55578..67242
/note="HURS-P3b repeat: matches 5213..6885 of consensus"
67618..70203
/note="LIM1 repeat: matches 858..3409 of consensus"
76561..76610
/note="25 copies 2 mer tt 78% conserved"
complement(77362..77634)
/note="match: GSS: Em:AQ731675"
77643..78226
/note="match: GSS: Em:AQ489768"
77671..78133
/note="match: GSS: Em:AQ285114"
78300..78420
/note="LIP3 repeat: matches 6024..6144 of consensus"
79742..80106
/note="LIP7 repeat: matches 5777..6141 of consensus"
82916..83206
/note="LIP15 repeat: matches 5838..6149 of consensus"
85946..87158
/note="LIM7 repeat: matches 4936..6168 of consensus"
87490..88402

Query Match      14.7%; Score 183.2; DB 9; Length 125133;
Best Local Similarity 86.0%; Pred. No. 1.6e-35;
Matches 203; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 694 CAGGAGCCATGTCGAATCTCAAAATTTGAAGAATCTTTGTCCCAACACACACC 753
Db 100433 CAGGAGTCATGCTCCTGAGCTAAATTTAAAGAATCTTTGTCCCAACACATACCA 100492

QY 754 AAAGAAAATAAACAAGAGGAGGAGGATGAAATTTGGCGTCTACCAACCCCTCCAGTA 813
Db 100493 ATTAATAATAAACAAGAGGAGGAGGATGAAATTTGGCGTCTACCAACCCCTCCAGTA 100552

QY 814 GCAGAAACACCTGTACCATCTCTTCAGTAACAGAAATAGAGACCCCACTGCAAGAATT 873
Db 100553 GCAGAAACATCTGTACCAACCTCTTCAGTAGAGAAATAGAGACCCCAATACAGAATT 100612

QY 874 CCGCGACTGTACCATAGCTGGAGGCCCTTAGGACATTTGCATTTTCACTATTTTC 929
Db 100613 TTACGCTCTGCTATAGCTGGAGGCCCTTAGGACCTTTGCACATTTTCCCATTTTC 100668
```

RESULT 6

```
AC093249
LOCUS      185664 bp      DNA      linear      PRI 03-FEB-2002
DEFINITION Homo sapiens chromosome 16 clone RP11-146F11, complete sequence.
ACCESSION  AC093249
VERSION     AC093249.3  GI:18481999
KEYWORDS   HTG.
SOURCE      Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 185664)
            DOE Joint Genome Institute.
            Sequencing of Human Chromosome 16
            Unpublished
            DOE Joint Genome Institute.
            2 (bases 1 to 185664)
            Direct Submission
            DOE Joint Genome Institute.
            Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            3 (bases 1 to 185664)
            Direct Submission
            DOE Joint Genome Institute.
            Submitted (03-FEB-2002) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            On Feb 3, 2002 this sequence version replaced gi:15383791.
            Sequence Quality Assessment:
            This entry has been annotated with sequence quality
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than
            1 error in 10,000 bp.
            Base-by-base quality values are not generally visible from the
            GenBank flat file format but are available as part
            of this entry's ASN.1 file.
            -----
            Sequence Quality Assessment:
            This entry has been annotated with sequence quality
            estimates computed by the Phrap assembly program.
            All manually edited bases have been reduced to quality zero.
            Quality levels above 40 are expected to have less than
            1 error in 10,000 bp.
            Base-by-base quality values are not generally visible from the
            GenBank flat file format but are available as part
            of this entry's ASN.1 file.
            -----
            FEATURES                     Location/Qualifiers
            source                       1..185664
                                         /organism="Homo sapiens"
                                         /db_xref="taxon:9606"
                                         /chromosome="16"
                                         /clone="RP11-146F11"
BASE COUNT  45014 a 47979 c 46282 g 46389 t
ORIGIN
Query Match      14.0%; Score 173.8; DB 9; Length 185664;
Best Local Similarity 86.1%; Pred. No. 4.1e-33;
Matches 204; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 694 CAGGAGCCATGTCGAATCTCAAAATTTGAAGAATCTTTGTCCCAACACACACC 753
Db 83718 CAGGAGCCATGTCGAATCTCAAAATTTCAAAAGATCTTTGTCCCTCCACAGCGCC 83777

QY 754 AAAGAAAATAAACAAGAGGAGGAGGATGAAATTTGGCGTCTACCAACCCCTCCAGTA 813
Db 83778 ATTG-AAATAAACAAGAGGAGGAGGATGAAATTTGGCTGTACCGCCCTCCAGTA 83836

QY 814 GCAGAAACACCTGTACCATCTCTTCAGTAACAGAAATAGAGACCCCACTGCAAGAATT 873
Db 83837 GTGGACACACGTGTACTGCTCTCTTCAGTAACAGAAATAGAGACCCCAACAGAATT 83896

QY 874 CCGCGACTGTACCATAGCTGGAGGCCCTTAGGACATTTGCATTTTCACTATTTTC 930
Db 83897 CTACGCTCTGCTGCCATAGCTGGAGGCCCTTCGAACCTTTGCTTTTCTATTCT 83953
```

```

RESULT 7
AC106886
LOCUS
DEFINITION Homo sapiens chromosome 16 clone RP11-2C24, WORKING DRAFT SEQUENCE,
4 ordered pieces.
ACCESSION AC106886
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 207538)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 16
Unpublished
2. (bases 1 to 207538)
DOE Joint Genome Institute.
Direct Submission
Submitted (13-JAN-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3. (bases 1 to 207538)
DOE Joint Genome Institute.
Direct Submission
Submitted (27-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Mar 27, 2002 this sequence:
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 10835: contig of 10835 bp in length
* 10836 10935: gap of unknown length
* 10936 20026: contig of 9091 bp in length
* 20027 20126: gap of unknown length
* 20127 148471: contig of 128345 bp in length
* 148472 148571: gap of unknown length
* 148572 207538: contig of 58967 bp in length.
FEATURES
Location/Qualifiers
1. .207538
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-2C24"
BASE COUNT 52549 a 53103 c 52637 g 48945 t 304 others
ORIGIN

Query Match 14.0%; Score 173.8; DB 2; Length 207538;
Best Local Similarity 86.1%; Pred. No. 4 1e-33;
Matches 204; Conservative 0; Mismatches 32; Indels 1; Gaps 1;

QY 694 CAGGAGCCATGTTCTCAATCTCAAAATTTTGAAGAATCTTTTGTCCACCCACACACCC 753
Db 205252 CAGGAGCCATGTTCTCAATCTCAAAATTTTGAAGAATCTTTTGTCCCTCAACAGCGCC 205311

QY 754 AAAGAAAATAATAACAGGAGGAGGATGAAAATTTGGCGTCTAGCACCCCTCCAGTA 813
Db 205312 ATTG-AAATAATAACAGGAGGAGGAGTGAATTTGGCTGTATCCGCCCTCCAGTA 205370

QY 814 GCAGAAACACCTCTACCATCTCTCTTCAGTAACAGAAATAGAGACCCCACTGCAAGAAAT 873
Db 205371 GTGGACACACGTTACTGTCTCTTCAGTACAGAAATAGAGACCCCAACAAAGAAAT 205430

QY 874 CGCGGCACTGTACCATAGCTAGGAGAGCCCTTAGGACATTCAGCTTCACTATTTCT 930
Db 205431 CTACGCTCTGCTGCCATAGCTAGGAGAGCCCTTCGAACCTTGCTTTCTCTATTTCT 205487

RESULT 8
AL713966
LOCUS
DEFINITION Human DNA sequence from clone XXbac-161M6 on chromosome 6, complete
sequence.
ACCESSION AL713966
VERSION AL713966.7 GI:20196696
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Tracey, A.
Direct Submission
Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 18, 2002 this sequence version replaced gi:20088817.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep/xxbac-161M6 is
from a CHORI-501 human bac - PGF cell line library VECTOR:
PTARBAC2.1
This sequence was generated from part of bacterial clone contigs
constructed by the MHC Haplotype Consortium and collaborators.
Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC.
FEATURES
Location/Qualifiers
1. .124047
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="XXbac-161M6"
/clone_lib="CHORI-501"
BASE COUNT 35394 a 26727 c 24564 g 37362 t
ORIGIN

```



```
repeat_region 9482..10338
/note="L1M4 repeat: matches 3090..4067 of consensus"
repeat_region 10347..11410
/note="L1M2 repeat: matches 4268..5120 of consensus"
repeat_region 11411..11709
/note="AluY repeat: matches 1..295 of consensus"
repeat_region 11710..11974
/note="L1M2 repeat: matches 5120..5387 of consensus"
repeat_region 11975..12715
/note="L1M2 repeat: matches 40..1523 of consensus"
repeat_region 12724..13852
/note="HERV9 repeat: matches 7281..8399 of consensus"
repeat_region 13842..15313
/note="HERV9 repeat: matches 1730..3199 of consensus"
repeat_region 15316..16837
/note="HERV9 repeat: matches 1..1522 of consensus"
repeat_region 16838..17404
/note="L1M2 repeat: matches 1..560 of consensus"
repeat_region 17407..17957
/note="L1M2 repeat: matches 5567..6122 of consensus"
repeat_region 18022..18115
/note="L1M4 repeat: matches 7885..7977 of consensus"
repeat_region 18139..18317
/note="L1M2/D repeat: matches 5465..5657 of consensus"
repeat_region 18346..18601
/note="L1M2 repeat: matches 6057..6313 of consensus"
misc_feature 18406..18673
/note="match: GSS: Em:AQ388150"
repeat_region 18673..19646
/note="L1M5 repeat: matches 1..969 of consensus"
repeat_region 19879..20143
/note="MER51B repeat: matches 3..266 of consensus"
repeat_region 20146..20627
/note="L1M2 repeat: matches 5567..6057 of consensus"
repeat_region 20632..21122
/note="LOR1b repeat: matches 1..497 of consensus"
repeat_region 21126..21331
/note="L1M2 repeat: matches 6062..6260 of consensus"
repeat_region 21332..21639
/note="AluVa5 repeat: matches 1..308 of consensus"
misc_feature 21640..21934
/note="match: GSS: Em:AQ231450"
repeat_region 21640..21694
/note="L1M2 repeat: matches 6260..6315 of consensus"
repeat_region 21787..22733
/note="MER52A repeat: matches 270..1239 of consensus"
repeat_region 23452..23523
/note="36 copies 2 mer tt 75% conserved"
repeat_region 23525..23667
/note="AluDb repeat: matches 161..311 of consensus"
repeat_region 24292..24481
/note="MIR repeat: matches 51..243 of consensus"
repeat_region 24918..25221
/note="AluY repeat: matches 1..302 of consensus"
gene 26169..39233
/gene="HLA-DRB3"
mRNA join(26169..26304,33986..34255,36530..36811,37496..37606,
38080..38103,38903..39233)
/gene="HLA-DRB3"
/product="DJ172K2.6 (major histocompatibility complex,
class II, DR beta 3*01012)"
/note="match: CDNAS: Em:X03069 Em:E00199 Em:V00522
Em:A06800 Em:M77155 Em:M27635
match: ESTs: Em:AI750149 Em:AA443496 Em:W52776 Em:AA411478
Em:AA292976 Em:AA319703 Em:AA573731 Em:AA715935
Em:AI701955 Em:AI440139 Em:AA858441 Em:AI567982
Em:AI660611 Em:AI635249"
/evidence="not_experimental"
misc_feature join(26205..26304,33986..34255,36530..36811,37496..37606,
38080..38103,38903..38926)
/gene="HLA-DRB3"
/note="match: STS: Em:G28562"
CDS join(26205..26304,33986..34255,36530..36811,37496..37606,
```

```
38080..38103,38903..38916)
/gene="HLA-DRB3"
/note="match: proteins: Tr:P79483"
/codon_start=1
/evidence="not_experimental"
/product="DJ172K2.6 (major histocompatibility complex,
Query Match 13.9%; Score 173; DB 9; Length 131234;
Best Local Similarity 83.1%; Pred. No. 6.4e-33;
Matches 197; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 694 CAGGAGCCATGTTCTGAATCTCAAAATTTTGAAGAATCTTTTGTCCACCCACACACCC 753
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 56534 CAGGAGTGTATGTTCTGAGTCTAAAAATCTGAAAAATCTGTTGTCCACCCACAGCTCCA 56593
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 754 AAAGAAAATAAATAACAGGAGGAGGATGAAATTTGGCGTCTACCAACCCCTCCAGTA 813
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 56594 ATTGAAATATAAAACAGAGAGGAGGAGTAAATAATTTGGCTATACCACTCTCCAATA 56653
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 814 GCAGAAACACCTGTACCATCTCTTTCAGTAACAGAAATAGAGACCCCACTGCAAGAAAT 873
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 56654 GCAGAAACATCTGTACTGCTCTCTTCGGTAGCAGAAATAGAGACCCCACTGAGAAAT 56713
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 874 CGCGGACTCTACCATAGCTGGAGAGCCCTTAGGCATTTGACATTTTCACTATTTCT 930
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 56714 TTACACTCTGCTGCATAGCTGGAGAGCCCTTGAGACCTTTTCCCTATTTCT 56770
||||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AL670296/c 160623 bp DNA linear PRI 24-APR-2002
LOCUS Human DNA sequence from clone XxBac-213L12 on chromosome 6,
DEFINITION complete sequence.
ACCESSION AL670296
VERSION AL670296.7 GI:20135767
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
Tracey, A.
Direct Submission
Submitted (23-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 10, 2002 this sequence version replaced gi:19848110.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep XxBac-213L12 is
from a CHORI-502 human bac - Cox cell line library VECTOR:
PTARBAC2.1
This sequence was generated from part of bacterial clone contigs
constructed by the MHC Haplotype Consortium and collaborators.
Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC.
Location/Qualifiers
1..160623
```

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /clone="XXbac-213L12"
 /clone_lib="CHORI-502"
 BASE COUNT 47478 a 31913 c 32862 g 48370 t
 ORIGIN

Query Match 13.9%; Score 173; DB 9; Length 160623;
 Best Local Similarity 83.1%; Pred. No. 6.5e-33;
 Matches 197; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 694 CAGGAGCATGCTTCTGAATCTCAATTTTGAAGAATCTTTGTGCCACCCACACACCC 753
 Db 153404 CAGGAGTGATGCTTCTGAGTCTAAAAATCTGAAAAATCTGTGTGCCACCCACAGCTCCA 153345
 QY 754 AAGAAATATAACAGGAGGAGGAGGATGAAAAATTTGGCGTCTACACCCCTCCAGTA 813
 Db 153344 ATTGAAAAATAAAACAGGAGGAGGATGAAAAATTTGGCGTCTATACCACTCTCCCAATA 153285
 QY 814 GCAGAACACCTGTACCATCTCTCTAGTACAGAAATAGAGACCCCACTGCAAGAAAT 873
 Db 153284 GCAGAACATCTGTACTGTCTCTCTCTAGTACAGAAATAGAGACCCCAATACAGAAAT 153225
 QY 874 CCGCGGACTGCTACCATAGCTGGAGACCCCTTAGGACATTCGACTTTCATTTCT 930
 Db 153224 TTCACTCTGCTGCATAGCTGGAGACCCCTTAGGACATTCGACTTTCATTTCT 153168

RESULT 11
 AP000802/c 183165 bp DNA linear PRI 30-MAR-2001
 LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-629G13,
 DEFINITION complete sequences.
 ACCESSION AP000802
 VERSION AP000802.4 GI:13488765
 KEYWORDS HTG.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (sites)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Title
 Homo sapiens genomic DNA
 Published Only in DataBase (1999)
 REFERENCE 2 (bases 1 to 183165)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Title
 Direct Submission
 Submitted (30-NOV-1999) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 On Mar 29, 2001 this sequence version replaced gi:10130031.
 FEATURES
 Location/Qualifiers
 1..183165
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="RP11-629G13"

BASE COUNT 52120 a 36917 c 37962 g 56166 t
 ORIGIN

Query Match 13.6%; Score 169; DB 9; Length 183165;
 Best Local Similarity 84.4%; Pred. No. 6.8e-32;
 Matches 190; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 707 CTGAATCTCAATTTTGAAGAATCTTTGTGCCACCCACACACCAAGAAATAATA 766
 Db 63869 CTGAATCTCAATTTTGAAGAATCTTTGTGCCACCCACACAGCTCCCACTGAAAAATAAA 63810

QY 767 AACAGGAGGAGGAGTGAATAATTTGGGCTCTACCAACCCCTCCAGTAGCAGAAACACCTG 826
 Db 63809 AACAGGAGGAGGAGTGAATAATTTGGTCTATACGCCCTCCCTCCAGTTGCAGAAACATCTG 63750
 QY 827 TACCATCTCCTTCAGTACAGAAATAGAGACCCCACTGCAAGAAATTCGCGGACTGCTA 886
 Db 63749 TACCACCTCCTTCAGTACAGAAATAGAGACCCCACTGCAAGAAATTTATATCTGCTG 63690
 QY 887 CCATAGCTGGAGAGCCCTTAGGACATTCGACTTTCATCTATTCTC 931
 Db 63689 CCATAGCTGGAGAGCCCTTAGGACCTTGCACCTTTTCCTATTCTCC 63645

RESULT 12
 AC016902/c 184841 bp DNA linear PRI 09-MAY-2001
 LOCUS Homo sapiens BAC clone RP11-217B19 from 11, complete sequence.
 DEFINITION AC016902
 ACCESSION AC016902.4 GI:11038566
 VERSION HTG.
 KEYWORDS Homo sapiens.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 184841)
 Sulston,J.E. and Waterston,R.
 Title
 Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 184841)
 Grewal,N., Maupin,R., Gregory,S. and O'Rourke,K.
 Title
 The sequence of Homo sapiens BAC clone RP11-217B19
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 184841)
 Waterston,R.H.
 Title
 Direct Submission
 Submitted (08-DEC-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 184841)
 Waterston,R.H.
 Title
 Direct Submission
 Submitted (30-OCT-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 184841)
 Waterston,R.
 Title
 Direct Submission
 Submitted (09-MAY-2001) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Oct 30, 2000 this sequence version replaced gi:7630895.
 COMMENT
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics -----
 ----- Center project name: H_NH0217B19 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by

Query Match	13.6%	Score 169;	DB 2:	Length 185569;
Best Local Similarity	84.4%	Pred. No. 6.Be-32;		
Matches 190;	Conservative 0;	Mismatches 35;	Indels 0;	Gaps 0;
QY 707	CTGAATCTCAAAATTTTGAAGAATCTTTGTGCCACCCACACACCCAGTCCCAACTGAAATAATA	766		
Db 31855	CTGAATCTCAAAATTTTGAAGAATCTGTGTGCCACCCACACACCCAGTCCCAACTGAAATAATA	31796		
QY 767	AACAGGAGAGGAGGATGAAAAATTTGGCGTCTTACCACCCCTCCAGTAGCAGAAACACCTG	826		
Db 31795	AACAGGAGAGGAGGATGAAAAATTTGGCTATACCGCTCTCCAGTTCGACAAACATCTG	31736		
QY 827	TACCATCTCTTCTAGTACAGAAATAGAGACCCCAAGAAATTCGCGGACTGCTA	886		
Db 31735	TACCACCTCTCTAGTACAGAAATGAGACCCCAATACAAAGAATTTTACTCTGCTG	31676		
QY 887	CCATAGCTGGAGACCTCTAGGACATTCGACTTTTCACTATTTC	931		
Db 31675	CCATAGCTGGAGACCTCTAGGACCTTGACATTTTCTTATTTCCC	31631		
RESULT 14				
AF079797				
LOCUS	AF079797	7736 bp	DNA	linear
DEFINITION	Homo sapiens human endogenous retrovirus HML6 gag, pol, and env			
ACCESSION	pseudogenes, complete sequence.			
VERSION	AF079797			
KEYWORDS	AF079797.1 GI:4894422			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 7736)			
TITLE	Medstrand, P., Mager, D.L., Yin, H., Dietrich, U. and Blomberg, J. Structure and genomic organization of a novel human endogenous retrovirus family: HERV-K (HML-6)			
JOURNAL	J. Gen. Virol. 78 (Pt 7), 1731-1744 (1997)			
MEDLINE	93368436			
PUBMED	9225050			
REFERENCE	2 (bases 1 to 7736)			
AUTHORS	Yin, H., Medstrand, P., Kristofferson, A., Dietrich, U., Aman, P. and Blomberg, J.			
TITLE	Characterization of human MMV-like (HML) elements similar to a sequence that was highly expressed in a human breast cancer: further definition of the HML-6 group			
JOURNAL	Virology 256 (1), 22-35 (1999)			
MEDLINE	99189359			
PUBMED	10087223			
REFERENCE	3 (bases 1 to 7736)			
AUTHORS	Medstrand, P., Yin, H., Dietrich, U. and Blomberg, J.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-JUL-1998) Department of Medical Science, Section of Virology, Dag Hammarskjoldsvag 17, Uppsala 751 85, Sweden			
FEATURES	Location/Qualifiers			
source	1..7736			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
repeat_region	<1..7736			
LTR	/rpt_family="Human endogenous retrovirus HML6"			
gene	<1..687			
	688..2180			
	/gene="gag"			
	/note="putative"			
	/pseudo			
CDS	688..2180			
	/gene="gag"			
	/note="similar to the gag region of MMV and HERV-K retroviruses"			
	/pseudo			
gene	/codon_start=1			
	2181..5320			

Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.1.
SHGC-37302 G30664
Bases 17171 to 18499 excised as IS:Tn10.

FEATURES

source
1..35715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="LINLR-247E3"
BASE COUNT 9675 a 7687 c 8078 g 10275 t
ORIGIN

Query Match 13.6%; Score 168.8; DB 9; Length 35715;
Best Local Similarity 82.2%; Pred. No. 6.9e-32;
Matches 194; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 694 CAGGAGCCATGTTCTGAATCTCRAAATTTGAAGAATCTTTGTCCACCACACACCCC 753
Db 25727 CTGAATCATATTTCTGAGTCTAAAAATCTGAAGAATCTTTGTCCACCACACATACCA 25668
QY 754 AAAGAAAATAATAACAGAGAGGAGGATCAAAAATTTGGCGTCTTACCACCCCTCCAGTA 813
Db 25667 ACTGAAAATAATAACAGAGAGGAGGATCAAAAATTTGGCGTCTTACCACCCCTCCAGTA 25608
QY 814 GCAGAACACCTGTACCATCTCCTTACGTACAGAAATAGAGACCCCACTGCAAAAGAAAT 873
Db 25607 GCAGAAACATCTGTACTCTCCTTCCGTAGCAGAAATAGAGACCCCACTGCAAAACAACT 25548
QY 874 CCGGGAGTCTACCATAGCTGGAGAGCCCTTAGGACATTCACATTTTCATTTTC 929
Db 25547 TTATGCTCTGCTGCCATAGCTGGAGAGCCCTTAGGACCTTGGCGTTTTCCTATTTTC 25492

Search completed: December 24, 2002, 20:25:29
Job time : 3918.88 secs

THIS PAGE BLANK (uspto)